Monthly Forecasting of GDP with Mixed Frequency Multivariate Singular Spectrum Analysis

Hossein Hassani | António Rua | Emmanuel Sirimal Silva | Dimitrios Thomakos
Monthly Forecasting of GDP with Mixed Frequency Multivariate Singular Spectrum Analysis

Hossein Hassani | António Rua | Emmanuel Sirimal Silva | Dimitrios Thomakos
Abstract
The literature on mixed-frequency models is relatively recent and has found applications across economics and finance. The standard application in economics considers the use of (usually) monthly variables (e.g. industrial production) in predicting/fitting quarterly variables (e.g. real GDP). In this paper we propose a Multivariate Singular Spectrum Analysis (MSSA) based method for mixed frequency interpolation and forecasting, which can be used for any mixed frequency combination. The novelty of the proposed approach rests on the grounds of simplicity within the MSSA framework. We present our method using a combination of monthly and quarterly series and apply MSSA decomposition and reconstruction to obtain monthly estimates and forecasts for the quarterly series. Our empirical application shows that the suggested approach works well, as it offers forecasting improvements on a dataset of eleven developed countries over the last 50 years. The implications for mixed frequency modelling and forecasting, and useful extensions of this method, are also discussed.

JEL: C18, C53, E17
Keywords: multivariate SSA, mixed frequency, GDP, industrial production, forecasting.
1. Introduction

Over the past two decades there has been an expanding literature on the theoretical foundations and empirical applications of Singular Spectrum Analysis (SSA), (see for example Kalantari et al. (2018), Hassani et al. (2018), Silva et al. (2017), and references therein). As SSA allows to cope with both linear and nonlinear as well as with stationary and non-stationary time series structures it turns out to be a fairly versatile approach for modelling and forecasting (Sanei and Hassani, 2015). Although SSA is a powerful tool in time series analysis and has been applied in a wide range of fields, only recently there has been an increasing interest in the use of SSA in economics and finance (see, for example, Hassani and Thomakos (2010) and Hassani and Patterson (2014) for two comprehensive overviews).

The use of SSA based models has proved to be reasonably useful for forecasting economic and financial variables. For instance, SSA has been used to forecast economic activity namely industrial production (Hassani et al. (2009a, 2013, 2016), Patterson et al. (2011), Silva et al. (2018a)), and GDP (Hassani and Zhigljavsky (2009), Hassani et al. (2011)). See also Hassani et al. (2013) who consider a set of UK economic variables including GDP and industrial production whereas Papailias and Thomakos (2017) consider a set of US variables including GDP. Additionally, de Carvalho et al. (2012) and de Carvalho and Rua (2017) resort to SSA to nowcast the US output gap, a first application of the ideas for mixed frequency SSA. SSA has also been used for forecasting tourism, in particular UK tourism income by Beneki et al. (2012), US tourist arrivals in Hassani et al. (2015) and European tourist arrivals by Hassani et al. (2017). Applications exploiting SSA for forecasting exchange rates (Lisi and Medio (1997) and Hassani et al. (2009b)), inflation (Hassani et al. (2013), Silva et al. (2018b)), and energy (Beneki and Silva (2013), Silva (2013)) are also available.

In this paper, we attempt to solve the mixed-frequency interpolation/forecasting problem in the context of SSA. Such an approach has not been considered in the previous literature and has certain a priori advantages which suggests that we examine its efficacy: first, SSA is a complete, stand-alone smoothing/filtering/forecasting method; second, SSA is model-free and its performance has been found to be very good across different types of time series (Ghodsi et al. (2017)); and third, SSA can easily handle multivariate applications.

Early works on the use of mixed frequencies relate with the strand of literature that focus on temporal disaggregation of time series, namely by obtaining high frequency estimates of a series observed at a lower frequency resorting to high frequency indicators. Naturally, the underlying relationship can be used to obtain estimates for the out of sample period covered by the low frequency variable. In this respect, one should mention the well-known seminal work of Chow and Lin (1971) who developed a regression based framework
for temporal disaggregation. However, the usefulness of the proposed method was limited in practice as it depends on the validity of the regression model which is bound by a variety of parametric assumptions which are unlikely to hold in the real world. Subsequent work includes Fernandez (1981), Litterman (1983), Wei and Stram (1990), Guerrero (1990), Liu and Hall (2000) and Santos Silva and Cardoso (2001), among others. Recently, the mixed frequency models regained increased interest with, for example, the MIDAS (mixed-data sampling) approach of Eric Ghysels and co-authors (see, for example, Ghysels et al. (2007) and Andreou et al. (2010, 2013)). The body of work using the MIDAS approach for forecasting includes Kuzin et al. (2011), Clements and Galvão (2012), Monteforte and Moretti (2013), Galvão (2013), Ghysels and Ozkan (2015), Duarte et al. (2017), among others. In addition, dynamic factor models for mixed frequency forecasting have also been developed (see, for example, Schumacher and Breitung (2008), Mariano and Murasawa (2003), Foroni and Marcellino (2014), Marcellino et al. (2016)). Other mixed-frequency approaches have been pursued by, for example, Gotz et al. (2014), Carriero et al. (2015), Schorfheide and Song (2015), Foroni et al. (2015), Barsoum and Stankiewicz (2015) and Marcellino and Sivec (2016).

All the above approaches, and their refinements, not only attract much attention in the literature but also offer solutions for faster tracking of important economic variables, mainly GDP, by approximating their paths with the use of auxiliary, correlated variables that are observed at higher frequencies (such as monthly, weekly or even daily). Based on the above mentioned literature, it is evident that mixed frequency forecasting is an important topic. However, the majority of the models used for mixed frequency forecasting are restricted by their parametric nature and related assumptions pertaining to normality, linearity and stationarity. Herein, we extend the list of models used for mixed frequency forecasting by resorting to the Multivariate Singular Spectrum Analysis (MSSA). Whilst SSA and MSSA were initially used for nowcasting through the work done by de Carvalho and Rua (2017), this paper marks the introduction of MSSA for mixed frequency forecasting. Given its nonparametric nature, modelling with MSSA enables users to ensure there is no loss of information as data transformations are not required, and it enables the smoothing, filtering and signal extraction which can also be useful when modelling data with mixed frequencies.

We follow the same underlying intuition of the other methods in the literature in developing our approach in the simplest possible context: interpolating and forecasting the monthly path of GDP, which is only observed at the quarterly frequency, by using industrial production as a highly correlated monthly proxy. To highlight the usefulness of our suggested approach, based on mixed frequency multivariate SSA, we consider an empirical application using data for eleven developed countries for the period running from the beginning of 1960 up to the end of 2013. We find that once a monthly measure of GDP growth has been obtained through the mixed frequency multivariate
SSA, one can improve the forecasting performance substantially by taking into account the monthly dynamics vis-à-vis the case where one forecasts the quarterly series. The forecasting gains are noteworthy across all countries with the average improvement in terms of the Root Mean Squared Forecast Error (RMSFE) and Mean Absolute Forecast Error (MAFE) being around 40 per cent. We also find that, even in a pseudo real time environment, the suggested method still improves, although slightly, the quarterly counterpart. Moreover, one should stress that this approach allows to track on a monthly basis economic developments which is *per se* valuable for real time monitoring and policymaking.

The remainder of the paper is organized as follows. In Section 2 we introduce the novel approach based on multivariate SSA for coping with a mixed frequency data framework. In Section 3 the dataset considered is described. In Section 4, the empirical application is conducted and the results are discussed. Finally, Section 5 concludes.

2. A Mixed Frequency Multivariate SSA Approach

2.1. Multivariate SSA: Decomposition and Reconstruction

In this section, we review (one of the possible ways of doing) multivariate SSA. Our presentation is for the bivariate case but pure multivariate adaptations are straightforward. Therefore, consider the bivariate time series \( \{ \mathbf{X}_t \triangleq [X_{t1}, X_{t2}]^\top \} \) taking values in \( \mathcal{R}_X \subseteq \mathbb{R} \). The index set \( S \) can be either \( \mathbb{Z} \) or \( \mathbb{N} \), thus covering the case of stationary and nonstationary time series. It is assumed that both series are already appropriately scaled and expressed in commensurable units of measurement. This is important when we introduce mixed frequencies in the next sub-section.

Suppose that we have available a sample of size \( N \) and let \( m \) denote the embedding dimension we propose to use. Applying the hankelization operator \( \mathcal{H}_m(\cdot) \) to each of the component series of \( \mathbf{X}_i \) we obtain the trajectory \((n \times m)\) matrices \( \mathbf{T}_i \triangleq \mathcal{H}_m(X_{1i}, X_{2i}, \ldots, X_{Ni}) \), for \( i = 1, 2 \) and \( n = N - m + 1 \). Concatenating the trajectory matrices horizontally we obtain the \((n \times 2m)\) MSSA trajectory matrix \( \mathbf{T}_X \triangleq [\mathbf{T}_1, \mathbf{T}_2] \) which we use for decomposition and reconstruction.

The \((2m \times 2m)\) sample covariance matrix is then defined as:

\[
\mathbf{C} \triangleq n^{-1} \mathbf{T}_X^\top \mathbf{T}_X
\]

which is block-symmetric and given by:

\[
\mathbf{C} = \begin{bmatrix}
\mathbf{C}_{11} & \mathbf{C}_{12} \\
\mathbf{C}_{21} & \mathbf{C}_{22}
\end{bmatrix},
\]

where \( \mathbf{C}_{ij} \) are correlation matrices and \( \mathbf{C}_{12} = \mathbf{C}_{21}^\top \).
such that it contains the own and cross-covariance matrices as its elements. Denote the spectral decomposition of $C$, in standard notation, by:

$$C_n \overset{def}{=} V\Lambda V^\top = \sum_{j=1}^{2m} \lambda_j v_j v_j^\top$$

with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2m}$. Splitting the $(2m \times 2m)$ matrix $V$ that holds the eigenvectors appropriately to $V \overset{def}{=} [V_1^\top, V_2^\top]^\top$ we can estimate the individual trajectory matrices as:

$$\hat{T}_i(k) \overset{def}{=} T_i Q(k)$$

where $Q(k) \overset{def}{=} \sum_{j \in I_k} v_{ij} v_{ij}^\top$ and $V_i \overset{def}{=} [v_{i1}, v_{i2}, \ldots, v_{im}]$ for $i = 1, 2$. Here $k \overset{def}{=} \text{dim} I_k \leq m$ and $I_k$ denotes the set of eigenvalue indices used in the reconstruction process. Finally, applying the diagonal averaging operator $D_{(m,N)}(\cdot)$ to the estimated trajectory matrix as in:

$$\{\hat{X}_{ti}(k)\}_{t=1}^N \overset{def}{=} D_{(m,N)}[\hat{T}_i(k)]$$

we obtain the reconstructed series.

### 2.2. MSSA when one series has higher frequency

We now turn to the proposed methodology for MSSA when one of the series exhibit higher sampling frequency, i.e. mixed frequency MSSA or MFMSSA. To make things concrete, and to relate them to our empirical results, let us suppose that $X_{t1}$ is a quarterly time series and $X_{t2}$ is a monthly time series. Now $N$ denotes the length of the higher frequency series and note that we must have expressed the units of measurement of the two series so that they match the higher sampling frequency. For example, when dealing with a quarterly and a monthly series, which are expressed as growth rates, we must transform them so that their growth rates correspond to the same period and the same sampling interval. In particular, in our empirical application we take the quarterly growth rate of GDP to be the value observed at the third month of the quarter whereas the values for the remaining months of the quarter are unknown. In the case of industrial production, the value for each month corresponds to the quarter-on-quarter growth rate for the quarter ended at that month. Hence, as a first step we prepare our data according to the following format, and mark the positions
of the quarterly series for which actual data are available:

\[
\begin{pmatrix}
(Q1, M1) & X_{11} & X_{21} & 0 \\
(Q1, M2) & X_{11} & X_{22} & 0 \\
(Q1, M3) & X_{11} & X_{23} & 1 \\
(Q2, M4) & X_{21} & X_{24} & 0 \\
(Q2, M5) & X_{21} & X_{25} & 0 \\
(Q2, M6) & X_{21} & X_{26} & 1 \\
(Q3, M7) & X_{31} & X_{27} & 0 \\
(Q3, M8) & X_{31} & X_{28} & 0 \\
(Q3, M9) & X_{31} & X_{29} & 1 \\
(Q4, M10) & X_{41} & X_{2,10} & 0 \\
(Q4, M11) & X_{41} & X_{2,11} & 0 \\
(Q4, M12) & X_{41} & X_{2,12} & 1
\end{pmatrix}, \tag{6}
\]

so that we initially fill-in the monthly values of the quarterly series with the actual, end-of-quarter, values. Note that we mark the positions/dates of the quarterly series which should not be changed during our iterations below with ones in the last column. That is, for the positions for which actual data are available we must retain them as such, and approximate the data for the other monthly positions for which we have no information.\(^1\)

The idea we use for the monthly interpolation is very simple: pass the initial data to the standard MSSA approach described before and obtain the fitted values, re-insert the actual values on the fixed positions indicated by the 1’s in equation (6), measure the mean-squared deviation between one round of approximation and the next and terminate when an appropriate condition is met. If we denote by \(\hat{X}_{t1}^{(r)}\) the value of the slower frequency series at the \(r^{th}\) round/iteration of the above procedure, we can schematically illustrate the method as follows:

Step 0. Using the data formatting in equation (6), apply MSSA to the two series and obtain the higher frequency fitted values \(\hat{X}_{t1}^{(0)}\). Initialize the mean-squared measure as

\[
RMSE(0) \overset{\text{def}}{=} \sqrt{N^{-1} \sum_{t=1}^{N} \left[ \hat{X}_{t1}^{(0)} \right]^2}.
\]

Step 1. Substitute the actual values indicated by the positions of 1’s in equation (6) into \(\hat{X}_{t1}^{(0)}\) and apply MSSA again to obtain the updated higher frequency fitted values \(\hat{X}_{t1}^{(1)}\); substitute again the actual values as before into this

---

1. Herein, the objective is to find an unknown high frequency series whose last values are consistent with a known low frequency series. Alternatively, depending on the application in mind, one could have considered to be consistent with the first values, sums or averages.
new series and compute the new mean-squared measure as $RMSE(1) \overset{\text{def}}{=} \sqrt{N^{-1} \sum_{t=1}^{N} \left[ \hat{X}_{t1}^{(0)} - \hat{X}_{t1}^{(1)} \right]^2}.$

Step 2. If $RMSE(1) \leq RMSE(0)$ but $|RMSE(1) - RMSE(0)| > \varepsilon$, for some small predefined number $\varepsilon$, then iterate again by passing $\hat{X}_{t1}^{(1)}$ to another round of MSSA to obtain $\hat{X}_{t1}^{(r)}$ for $r > 1$.

Step 3. While $RMSE(r + 1) \leq RMSE(r)$ but $|RMSE(r + 1) - RMSE(r)| > \varepsilon$ continue along steps 1 and 2; if $RMSE(r + 1) \leq RMSE(r)$ and $|RMSE(r + 1) - RMSE(r)| \leq \varepsilon$ terminate the iterations and use $\hat{X}_{t1}^{(r+1)}$ as the final estimate of the higher frequency series.

Note that the actual, appropriately measured, quarterly series are always used in the positions with 1’s in equation (6). Thus, only the missing months are filled-in with the interpolated values. In all our experiments we set $\varepsilon$ to the range $[1e^{-5}, 5e^{-4}]$ and convergence of the above algorithm was achieved in an average of less than 20 iterations.

There are a number of practical issues that the reader will immediately ask about: how should one select the higher frequency series so as to achieve meaningful interpolation values for the lower frequency series? What kind of embedding dimensions works well for this procedure, assuming that the first question is answered? How many eigenvectors should one use in the reconstruction phase of the higher frequency fitted values of the lower frequency series? These questions must be answered before one can implement the method. Here are our suggestions.

It is clear that the higher frequency series should be highly correlated if there is to be any a priori reason to believe that a good interpolation will take place. Therefore, we first seek a good (i.e. highly correlated) higher frequency proxies for the lower frequency series. If we agree on this suggestion, then one can consider doing a direct search on the length of the embedding dimension $m$ that maximizes the statistical correlation between the actual higher frequency series $X_{t2}$ and the fitted values of the lower frequency series $\hat{X}_{t1}^{(r)}$. That is, a plausible value $m^*$ for $m$ for our interpolation procedure might be selected to satisfy the condition:

$$m^* \overset{\text{def}}{=} \arg\max_m \text{Corr} \left( X_{t2}, \hat{X}_{t1}^{(r)} \right).$$  (7)

In our empirical application below, we found that selecting a value of $m$ equal to the higher sampling frequency (i.e. $m = 12$) was a reasonable compromise as it gave results practically identical to a search for $m^*$ as above. In essence, we found that the correlation was maximized at or close by to the higher sampling frequency. Such results merit further investigation with other kinds of series as well.

Finally, to answer the last question, we again looked at the ex-post correlation between $X_{t2}$ and $\hat{X}_{t1}^{(r)}$ and found that in the reconstruction phase
of MSSA essentially all eigenvectors should be used, that is all available statistical information from the decomposition of the joint trajectory matrix. This makes intuitive sense as there is no reason to believe that one can discard information from the higher frequency in obtaining interpolated values for the lower frequency and thus we consistently found that $k$ should be set equal to $m$.

3. Data

The data used in our empirical application include quarterly real GDP from Q1:1960 up until Q4:2013 and monthly Industrial Production from January 1960 up until December 2013 for 11 developed countries namely Austria, Belgium, France, Denmark, Italy, Japan, Netherlands, Portugal, Sweden, United Kingdom and United States. The data has been retrieved from the OECD Main Economic Indicators database. The data is referred to as mixed frequency owing to the GDP figures being quarterly whilst the IP figures are reported on a monthly basis. In general, the GDP growth and IP appears to be highly correlated for all countries and in this study we exploit this dependence for improving the forecastability of GDP. In Table 1, we provide a set of descriptive statistics regarding GDP and IP growth over the sample period considered.

Our analysis begins by considering quarterly GDP growth. The minimum and maximum columns indicate that during this period, the worst and best quarterly growth rates to be achieved during a particular quarter have been reported by France. In terms of the average quarterly GDP growth, the highest average has been reported by Japan whilst the lowest average quarterly GDP growth rate is equivalent in Germany and UK. Based on the standard deviation, the most variable quarterly GDP growth has been recorded in Netherlands with the most stable quarterly growth rate for GDP being recorded in US. The KS test for normality indicates evidence of normally distributed quarterly GDP growth six of these countries whilst all quarterly GDP growth series are found to be stationary based on an Augmented Dickey-Fuller (ADF) test at a $p$-value of 0.05. All correlations between GDP and IP growth were found to be statistically significant at the 5% significance level.

In terms of IP, we can perform a similar analysis. The worst and best monthly IP growth rates during this time period has been reported by Japan (worst) and France (best). Japan also has the highest average monthly IP growth rate whilst UK reports the lowest average growth. The KS test for normality indicates majority of the monthly IP series are skewed, and therefore if we give prominence to median growth as opposed to mean growth, the results still remain unchanged with the highest median monthly IP growth being reported in Japan with the lowest in UK. The SD criterion suggests Japan has the least stable average monthly IP growth rate and that US has
the most stable. If we consider the IQR as most data are skewed, then the least stable average IP growth is recorded in Portugal whilst US continues to report the most stable average IP growth rate. As with quarterly GDP growth rates there is no evidence in monthly IP growth rates for the time series to be nonstationary based on the ADF test for unit root problems. The coefficient of variation (CV) criterion enables us to compare the variation between the quarterly GDP growth and monthly IP growth. The CV clearly indicates that for all countries the monthly IP growth rate is more variable than the quarterly GDP growth.

We then go a step further and perform an ANOVA test to determine whether there are statistically significant differences between countries in terms of the quarterly GDP growth rates and monthly IP growth rates. Interestingly, based on the post-hoc Tukey HSD ($p = 0.05$) test we find evidence of statistically significant differences between the average quarterly GDP growth rates of only the following combination of countries, i.e. Germany and Japan, Italy and Japan, and Japan and UK. On the other hand, the post-hoc Tukey HSD test ($p = 0.05$) for statistically significant differences in the average monthly growth rates of IP had the following significant combinations. Accordingly, Austria and France, Austria and UK, Belgium and UK, France and Japan, Germany and Japan, Italy and Japan, Japan and Sweden, Japan and UK, Netherlands and UK, Portugal and UK, and UK and US were found

### Quarterly GDP

<table>
<thead>
<tr>
<th>Country</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Med</th>
<th>IQR</th>
<th>SD</th>
<th>CV</th>
<th>Skew</th>
<th>KS (p)</th>
<th>ADF</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>-4.38</td>
<td>4.51</td>
<td>0.72</td>
<td>0.73</td>
<td>1.16</td>
<td>1.04</td>
<td>144.93</td>
<td>-0.36</td>
<td>0.07</td>
<td>-5.97</td>
<td>0.52</td>
</tr>
<tr>
<td>Belgium</td>
<td>-2.16</td>
<td>3.76</td>
<td>0.67</td>
<td>0.73</td>
<td>0.89</td>
<td>0.74</td>
<td>110.42</td>
<td>-0.36</td>
<td>0.02</td>
<td>-4.58</td>
<td>0.49</td>
</tr>
<tr>
<td>Germany</td>
<td>-4.48</td>
<td>4.48</td>
<td>0.62</td>
<td>0.66</td>
<td>1.26</td>
<td>1.12</td>
<td>180.67</td>
<td>-0.34</td>
<td>0.03</td>
<td>-7.45</td>
<td>0.66</td>
</tr>
<tr>
<td>Italy</td>
<td>-2.88</td>
<td>6.00</td>
<td>0.63</td>
<td>0.56</td>
<td>1.29</td>
<td>1.04</td>
<td>166.46</td>
<td>0.53</td>
<td>0.03</td>
<td>-5.28</td>
<td>0.75</td>
</tr>
<tr>
<td>Japan</td>
<td>-3.97</td>
<td>5.7</td>
<td>0.99</td>
<td>0.91</td>
<td>1.60</td>
<td>1.34</td>
<td>135.12</td>
<td>-0.08</td>
<td>0.04</td>
<td>-4.50</td>
<td>0.63</td>
</tr>
<tr>
<td>Netherlands</td>
<td>-6.27</td>
<td>8.94</td>
<td>0.69</td>
<td>0.76</td>
<td>1.13</td>
<td>1.45</td>
<td>209.24</td>
<td>0.27</td>
<td>&lt;0.01</td>
<td>-8.14</td>
<td>0.38</td>
</tr>
<tr>
<td>Portugal</td>
<td>-2.33</td>
<td>4.90</td>
<td>0.81</td>
<td>0.81</td>
<td>1.46</td>
<td>1.22</td>
<td>151.38</td>
<td>0.16</td>
<td>0.20</td>
<td>-4.31</td>
<td>0.58</td>
</tr>
<tr>
<td>Sweden</td>
<td>-4.77</td>
<td>5.50</td>
<td>0.64</td>
<td>0.66</td>
<td>1.33</td>
<td>1.31</td>
<td>204.78</td>
<td>-0.18</td>
<td>&lt;0.01</td>
<td>-7.48</td>
<td>0.63</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>-2.41</td>
<td>5.28</td>
<td>0.62</td>
<td>0.65</td>
<td>0.84</td>
<td>0.97</td>
<td>157.89</td>
<td>0.43</td>
<td>&lt;0.01</td>
<td>-6.01</td>
<td>0.67</td>
</tr>
<tr>
<td>United States</td>
<td>-2.11</td>
<td>3.89</td>
<td>0.76</td>
<td>0.77</td>
<td>0.82</td>
<td>0.85</td>
<td>110.01</td>
<td>-0.27</td>
<td>&lt;0.01</td>
<td>-4.34</td>
<td>0.76</td>
</tr>
</tbody>
</table>

### Monthly IP

<table>
<thead>
<tr>
<th>Country</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Med</th>
<th>IQR</th>
<th>SD</th>
<th>CV</th>
<th>Skew</th>
<th>KS (p)</th>
<th>ADF</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>-6.44</td>
<td>7.22</td>
<td>0.92</td>
<td>1.04</td>
<td>2.23</td>
<td>1.75</td>
<td>189.93</td>
<td>-0.39</td>
<td>0.03</td>
<td>-9.71</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>-10.36</td>
<td>17.16</td>
<td>0.68</td>
<td>0.75</td>
<td>2.12</td>
<td>2.20</td>
<td>323.50</td>
<td>0.22</td>
<td>&lt;0.01</td>
<td>-13.87</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>-11.12</td>
<td>21.73</td>
<td>0.47</td>
<td>0.45</td>
<td>1.85</td>
<td>2.27</td>
<td>486.28</td>
<td>1.36</td>
<td>&lt;0.01</td>
<td>-15.99</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>-13.40</td>
<td>5.89</td>
<td>0.61</td>
<td>0.76</td>
<td>2.27</td>
<td>2.00</td>
<td>325.63</td>
<td>-1.73</td>
<td>&lt;0.01</td>
<td>-9.62</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>-10.85</td>
<td>15.24</td>
<td>0.54</td>
<td>0.56</td>
<td>2.49</td>
<td>2.47</td>
<td>459.33</td>
<td>-0.10</td>
<td>&lt;0.01</td>
<td>-11.44</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>-19.70</td>
<td>10.42</td>
<td>1.01</td>
<td>1.13</td>
<td>2.48</td>
<td>2.81</td>
<td>277.49</td>
<td>-1.99</td>
<td>&lt;0.01</td>
<td>-9.53</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>-7.30</td>
<td>9.82</td>
<td>0.70</td>
<td>0.77</td>
<td>2.09</td>
<td>2.01</td>
<td>285.62</td>
<td>-0.25</td>
<td>0.09</td>
<td>-12.20</td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>-7.51</td>
<td>10.86</td>
<td>0.77</td>
<td>0.70</td>
<td>2.83</td>
<td>2.49</td>
<td>325.5</td>
<td>0.41</td>
<td>&lt;0.01</td>
<td>-13.18</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>-11.47</td>
<td>7.16</td>
<td>0.63</td>
<td>0.64</td>
<td>2.40</td>
<td>2.19</td>
<td>356.96</td>
<td>-0.79</td>
<td>&lt;0.01</td>
<td>-12.23</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>-7.04</td>
<td>8.64</td>
<td>0.26</td>
<td>0.30</td>
<td>1.53</td>
<td>1.60</td>
<td>610.21</td>
<td>-0.03</td>
<td>&lt;0.01</td>
<td>-13.61</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>-7.38</td>
<td>4.51</td>
<td>0.69</td>
<td>0.90</td>
<td>1.41</td>
<td>1.58</td>
<td>228.36</td>
<td>1.27</td>
<td>&lt;0.01</td>
<td>-9.35</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* † indicates data is normally distributed based on the Kolmogorov-Smirnov (KS) test at a p-value of 0.01. r indicates the Pearsin correlation between GDP and IP. The * indicates a statistically significant correlation between GDP and IP quarterly growth at the usual significance level of 0.05.

Table 1. Descriptive statistics for GDP and IP growth rates.
to be countries with statistically significant differences in monthly IP growth rates.

4. Evaluating the MFMSSA approach

As motivated in Section 2.2, herein we conduct the following exercise. Based on the whole sample period, we take the quarterly GDP growth and perform the MFMSSA interpolation using as covariate the monthly industrial production growth. The resulting monthly estimates for GDP growth are displayed in Figure 1 alongside the corresponding covariate series. From Figure 1, one can see that the monthly estimates of GDP are much less volatile than industrial production and in Table 2 we report the corresponding correlations.

The weighted correlation (w-correlation) is another measure of dependence between two series which is considered to evaluate the appropriateness of the separability between signal and noise as achieved via MSSA. If we consider two time series $Y^{(1)}_N$ and $Y^{(2)}_N$, then the w-correlation can be calculated as:

$$\rho_{12}^{(w)} = \frac{\langle Y^{(1)}_N, Y^{(2)}_N \rangle_w}{\|Y^{(1)}_N\|_w \|Y^{(2)}_N\|_w},$$

where $Y^{(1)}_N$ and $Y^{(2)}_N$ are two time series, $\|Y^{(i)}_N\|_w = \sqrt{\langle Y^{(i)}_N, Y^{(i)}_N \rangle_w}$, $\langle Y^{(i)}_N, Y^{(j)}_N \rangle_w = \sum_{k=1}^N w_k y^{(i)}_k y^{(j)}_k$ ($i, j = 1, 2$), $w_k = \min\{k, L, N - k\}$ (here, assume $L \leq N/2$).

As explained in Hassani et al. (2012), if the absolute value of the w-correlations is small, then the corresponding series are almost w-orthogonal, but, if it is large, then the two series are far from being w-orthogonal and are therefore badly separable. As an example, we present in Figure 2 the w-correlation matrix for the United States to show the dependence between signal and noise components.

---

2. We also find that the IP series does not Granger cause the monthly GDP series, which seems natural as the information from the IP is already embedded in the estimation of the monthly GDP, but the monthly GDP series seems to Granger cause the IP series in most countries.
Figure 1: Monthly GDP and IP growth
Figure 1: Monthly GDP and IP growth (continued)
In Table 2, we also report the correlations of the MFMSSA estimates with those obtained with the widely used Chow-Lin method to disaggregate time series. The correlations range from 0.62 for the Netherlands up to 0.97 for Belgium attaining, on average across all countries, more than 0.8. As the Chow-Lin method, in its original formulation, is only applicable to static models, we also consider the dynamic model variant of the Chow-Lin method (see Santos Silva and Cardoso (2001)) as it is a more natural working assumption when dealing with time series. In this respect, the MFMSSA estimates present an even higher correlation for most countries. The correlations range from 0.72 for Sweden up to 0.99 for Belgium attaining, on average across all countries, a value close to 0.9. Such a comparison reinforces the plausibility of the monthly GDP estimates obtained with the MFMSSA approach.

<table>
<thead>
<tr>
<th>Country</th>
<th>IP</th>
<th>Chow-Lin</th>
<th>Dynamic model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>0.48</td>
<td>0.69</td>
<td>0.90</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.46</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>France</td>
<td>0.80</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>Germany</td>
<td>0.64</td>
<td>0.78</td>
<td>0.94</td>
</tr>
<tr>
<td>Italy</td>
<td>0.76</td>
<td>0.86</td>
<td>0.97</td>
</tr>
<tr>
<td>Japan</td>
<td>0.64</td>
<td>0.93</td>
<td>0.78</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.32</td>
<td>0.62</td>
<td>0.83</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.47</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.55</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.63</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>United States</td>
<td>0.76</td>
<td>0.85</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 2. Correlation of MFMSSA monthly GDP with monthly IP, monthly GDP with Chow-Lin method and monthly GDP with the dynamic extension of Chow-Lin method.
Another evaluation exercise of the MFMSSA approach consists in assessing its usefulness in a forecasting context. In this respect, one can compare the 3-step ahead monthly forecast of the interpolated series with the 1-step ahead quarterly forecast of the actual series. In particular, we take the full sample of observations, perform the MFMSSA interpolation and then take the resulting monthly estimates for GDP growth and forecast it 3-months ahead via a parsimonious AR model with the order length chosen according to standard information criteria. The forecasts obtained for the third month of each quarter can be directly compared with the ones obtained by simply fitting an AR model to the quarterly GDP series and forecasting 1-quarter ahead. This exercise allows us to assess how much is gained in terms of forecasting performance by taking on board the monthly dynamics obtained via MFMSSA.\footnote{As the quarterly GDP growth figures correspond to the values in the third month of the corresponding quarter in the monthly series as described in Section 2.2, when modelling the monthly series one is already mixing the quarterly with the monthly information in a similar spirit to the MIDAS approach. However, since a higher frequency series is being considered when fitting the model, a richer dynamics can be exploited. In fact, we find that resorting to a MIDAS approach does not deliver better forecasting results.}

We conduct the forecasting exercise using an expanding window and recursive model selection and estimation.\footnote{We also conducted the forecasting exercise using a rolling window scheme but the forecasting performance deteriorates. This may reflect the fact that with an expanding window more information is taken on board as time goes by.} Since the business cycle frequency range is typically defined in the literature to be between two and eight years, we considered a starting period of sixteen years so as to encompass at least two complete business cycles. The forecasting results of such an exercise are presented in Table 3. In particular, we report the relative RMSFE, i.e the ratio between the RMSFE of the monthly model and the RMSFE of the quarterly counterpart as well as the relative MAFE defined in a similar way. Furthermore, we computed the Superior Predictive Ability (SPA) test proposed by Hansen (2005) to compare the performance of the two forecasting models. Both loss functions, the mean squared error and the mean absolute error are considered and therefore we report $SPA_{MSE}$ and $SPA_{MAE}$ respectively. The statistic reported in the table refers to the $SPA$ p-value where a low value signals that the benchmark is outperformed by the MFMSSA based approach.

Table 3 shows that taking on board the monthly dynamics of GDP estimates can clearly improve the forecasting performance. In fact, both the RMSFE and the MAFE are lower in the case where the monthly model is used for forecasting 1-quarter ahead. This finding holds for all the countries under study and the gains are quite noteworthy across countries. The average improvement in terms of the RMSFE and MAFE is around 40 per cent. The countries where the gains are higher are Germany, Japan, Portugal, United Kingdom and United States whereas among the countries which present lower gains one should mention Belgium and Italy. Based on the $SPA$ test results, one can conclude that the
Table 3. Forecast evaluation with full sample estimates

<table>
<thead>
<tr>
<th>Country</th>
<th>RMSFE</th>
<th>MAFE</th>
<th>SPA_MSE</th>
<th>SPA_MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>0.606</td>
<td>0.613</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.728</td>
<td>0.825</td>
<td>0.049</td>
<td>0.033</td>
</tr>
<tr>
<td>France</td>
<td>0.615</td>
<td>0.619</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Germany</td>
<td>0.460</td>
<td>0.495</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Italy</td>
<td>0.822</td>
<td>0.792</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Japan</td>
<td>0.482</td>
<td>0.484</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.560</td>
<td>0.594</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.446</td>
<td>0.446</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.579</td>
<td>0.592</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.545</td>
<td>0.539</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>United States</td>
<td>0.542</td>
<td>0.528</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

MFMSA based approach outperforms the benchmark, at the usual statistical significance level, for all countries and for both loss functions.

To assess what would have been the forecasting performance in a pseudo real time scenario, we use the same recursive approach as before, perform the MFMSA interpolation within such a time window span and then forecast 3-months ahead in the case of the monthly model and 1-quarter ahead based on the quarterly model. The results are displayed in Table 4.

Table 4. Forecast evaluation with recursive sample estimates

<table>
<thead>
<tr>
<th>Country</th>
<th>RMSFE</th>
<th>MAFE</th>
<th>SPA_MSE</th>
<th>SPA_MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>0.996</td>
<td>0.996</td>
<td>0.435</td>
<td>0.446</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.881</td>
<td>0.949</td>
<td>0.139</td>
<td>0.222</td>
</tr>
<tr>
<td>France</td>
<td>0.803</td>
<td>0.777</td>
<td>0.008</td>
<td>0.001</td>
</tr>
<tr>
<td>Germany</td>
<td>0.944</td>
<td>0.958</td>
<td>0.070</td>
<td>0.129</td>
</tr>
<tr>
<td>Italy</td>
<td>0.925</td>
<td>0.900</td>
<td>0.056</td>
<td>0.021</td>
</tr>
<tr>
<td>Japan</td>
<td>0.986</td>
<td>0.981</td>
<td>0.365</td>
<td>0.318</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.934</td>
<td>0.985</td>
<td>0.069</td>
<td>0.346</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.975</td>
<td>0.973</td>
<td>0.128</td>
<td>0.162</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.972</td>
<td>1.014</td>
<td>0.256</td>
<td>0.599</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.957</td>
<td>0.995</td>
<td>0.159</td>
<td>0.440</td>
</tr>
<tr>
<td>United States</td>
<td>0.994</td>
<td>1.002</td>
<td>0.426</td>
<td>0.513</td>
</tr>
</tbody>
</table>

Table 4 shows that even when one takes into account the pseudo real time computation of the GDP monthly estimates, one still improves the 1-quarter ahead forecasts. However, in this case, the gains are much smaller being on average around 5 per cent. One notable case is France with the forecasting gains attaining around 20 per cent. Naturally, the SPA test results are weaker although in a few cases there is statistical evidence of forecasting improvement over the benchmark. Furthermore, one should stress that the computation of monthly estimates enables one to compute forecasts every month and therefore
one is not restricted to forecast only on quarterly basis. This can be a valuable feature for tracking economic conditions in a real life environment.

5. Conclusions

In this paper, we suggest a new approach based on SSA to cope with data sampled at different frequencies. In particular, we layout the theoretical foundations and rationale underlying such a mixed frequency multivariate SSA approach and address some practical issues related with its implementation.

In the empirical application, we consider two variables, namely GDP and industrial production. The former is sampled at the quarterly frequency whereas the latter is monthly. Resorting to the MFMSSA approach we obtain monthly estimates for GDP growth and assess the forecasting performance of a monthly model over the quarterly counterpart. We analyze a set of eleven developed countries over the period running from the beginning of 1960 up to the end of 2013.

The results obtained are quite promising. We find that taking on board the monthly dynamics obtained via the MFMSSA method allows to achieve noteworthy forecasting gains for all countries. Although the gains are lower in a pseudo real-time exercise, one should note that such an approach also enables one to deliver monthly forecasts which can constitute a valuable feature for monitoring economic evolution in a real life environment.
References


Working Papers

2017

1|17 The diffusion of knowledge via managers' mobility
Giordano Mion | Luca David Opromolla | Alessandro Sforza

2|17 Upward nominal wage rigidity
Paulo Guimarães | Fernando Martins | Pedro Portugal

3|17 Zooming the ins and outs of the U.S. unemployment
Pedro Portugal | António Rua

4|17 Labor market imperfections and the firm's wage setting policy
Sónia Félix | Pedro Portugal

5|17 International banking and cross-border effects of regulation: lessons from Portugal
Diana Bonfim | Sónia Costa

6|17 Disentangling the channels from birthdate to educational attainment
Luís Martins | Manuel Coutinho Pereira

7|17 Who's who in global value chains? A weighted network approach
João Amador | Sónia Cabral | Rossana Mastrandrea | Franco Ruzzenenti

8|17 Lending relationships and the real economy: evidence in the context of the euro area sovereign debt crisis
Luciana Barbosa

9|17 Impact of uncertainty measures on the Portuguese economy
Cristina Manteu | Sara Serra

10|17 Modelling currency demand in a small open economy within a monetary union
António Rua

11|17 Boom, slump, sudden stops, recovery, and policy options. Portugal and the Euro
Olivier Blanchard | Pedro Portugal

12|17 Inefficiency distribution of the European Banking System
João Oliveira

13|17 Banks' liquidity management and systemic risk
Luca G. Deidda | Ettore Panetti

14|17 Entrepreneurial risk and diversification through trade
Federico Esposito

15|17 The portuguese post-2008 period: a narrative from an estimated DSGE model
Paulo Júlio | José R. Maria

16|17 A theory of government bailouts in a heterogeneous banking system
Filomena Garcia | Ettore Panetti

17|17 Goods and factor market integration: a quantitative assessment of the EU enlargement
Florenzo Caliendo | Luca David Opromolla | Fernando Parro | Alessandro Sforza
2018

1|18  Calibration and the estimation of macro-economic models
   Nikolay Iskrev

2|18  Are asset price data informative about news shocks? A DSGE perspective
   Nikolay Iskrev

3|18  Sub-optimality of the friedman rule with distorting taxes
   Bernardino Adão | André C. Silva

4|18  The effect of firm cash holdings on monetary policy
   Bernardino Adão | André C. Silva

5|18  The returns to schooling unveiled
   Ana Rute Cardoso | Paulo Guimarães | Pedro Portugal | Hugo Reis

6|18  Real effects of financial distress: the role of heterogeneity
   Francisco Buera | Sudipto Karmakar

7|18  Did recent reforms facilitate EU labour market adjustment? Firm level evidence
   Mario Izquierdo | Theodora Kosma | Ana Lamo | Fernando Martins | Simon Savsek

8|18  Flexible wage components as a source of wage adaptability to shocks: evidence from European firms, 2010–2013
   Jan Babecký | Clémence Berson | Ludmila Fadejeva | Ana Lamo | Petra Marotzke | Fernando Martins | Pawel Strzelecki

9|18  The effects of official and unofficial information on tax compliance
   Filomena Garcia | Luca David Opromolla | Andrea Vezulli | Rafael Marques

10|18  International trade in services: evidence for portuguese firms
    João Amador | Sónia Cabral | Birgitte Ringstad

11|18  Fear the walking dead: zombie firms, spillovers and exit barriers
    Ana Fontoura Gouveia | Christian Osterhold

12|18  Collateral Damage? Labour Market Effects of Competing with China – at Home and Abroad
    Sónia Cabral | Pedro S. Martins | João Pereira dos Santos | Mariana Tavares

13|18  An integrated financial amplifier: The role of defaulted loans and occasionally binding constraints in output fluctuations
    Paulo Júlio | José R. Maria

14|18  Structural Changes in the Duration of Bull Markets and Business Cycle Dynamics
    João Cruz | João Nicolau | Paulo M.M. Rodrigues

15|18  Cross-border spillovers of monetary policy: what changes during a financial crisis?
    Luciana Barbosa | Diana Bonfim | Sónia Costa | Mary Everett

16|18  When losses turn into loans: the cost of undercapitalized banks
    Laura Blattner | Luísa Farinha | Francisca Rebelo

17|18  Testing the fractionally integrated hypothesis using M estimation: With an application to stock market volatility
    Matei Demetrescu | Paulo M. M. Rodrigues | Antonio Rubia
18 | 18 Every cloud has a silver lining: Micro-level evidence on the cleansing effects of the Portuguese financial crisis
Daniel A. Dias | Carlos Robalo Marques

19 | 18 To ask or not to ask? Collateral versus screening in lending relationships
Hans Degryse | Artashes Karapetyan | Sudipto Karmakar

20 | 18 Thirty years of economic growth in Africa
João Amador | António R. dos Santos

21 | 18 CEO performance in severe crises: the role of newcomers
Sharmin Sazedj | João Amador | José Tavares

22 | 18 A general equilibrium theory of occupational choice under optimistic beliefs about entrepreneurial ability
Michele Dell’Era | Luca David Opromolla | Luís Santos-Pinto

23 | 18 Exploring the implications of different loan-to-value macroprudential policy designs
Rita Basto | Sandra Gomes | Diana Lima

24 | 18 Bank shocks and firm performance: new evidence from the sovereign debt crisis
Luísa Farinha | Marina-Eliza Spaliara | Serafim Tsoukas

25 | 18 Bank credit allocation and productivity: stylised facts for Portugal
Nuno Azevedo | Márcio Mateus | Álvaro Pina

26 | 18 Does domestic demand matter for firms’ exports?
Paulo Soares Esteves | Miguel Portela | António Rua

27 | 18 Credit Subsidies
Isabel Correia | Fiorella De Fiore | Pedro Teles | Oreste Tristani
<table>
<thead>
<tr>
<th>Issue</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The transmission of unconventional monetary policy to bank credit supply: evidence from the TLTRO</td>
<td>António Afonso</td>
</tr>
<tr>
<td>2</td>
<td>How responsive are wages to demand within the firm? Evidence from idiosyncratic export demand shocks</td>
<td>Andrew Garin</td>
</tr>
<tr>
<td>3</td>
<td>Vocational high school graduate wage gap: the role of cognitive skills and firms</td>
<td>Joop Hartog</td>
</tr>
<tr>
<td>4</td>
<td>What is the Impact of Increased Business Competition?</td>
<td>Sónia Félix</td>
</tr>
<tr>
<td>5</td>
<td>Modelling the Demand for Euro Banknotes</td>
<td>António Rua</td>
</tr>
<tr>
<td>7</td>
<td>The new ESCB methodology for the calculation of cyclically adjusted budget balances: an application to the Portuguese case</td>
<td>Cláudia Braz</td>
</tr>
<tr>
<td>8</td>
<td>Into the heterogeneities in the Portuguese labour market: an empirical assessment</td>
<td>Fernando Martins</td>
</tr>
<tr>
<td>9</td>
<td>A reexamination of inflation persistence dynamics in OECD countries: A new approach</td>
<td>Gabriel Zsurkis</td>
</tr>
<tr>
<td>10</td>
<td>Euro area fiscal policy changes: stylised features of the past two decades</td>
<td>Cláudia Braz</td>
</tr>
<tr>
<td>11</td>
<td>The Neutrality of Nominal Rates: How Long is the Long Run?</td>
<td>João Valle e Azevedo</td>
</tr>
<tr>
<td>13</td>
<td>Monthly Forecasting of GDP with Mixed Frequency Multivariate Singular Spectrum Analysis</td>
<td>Hossein Hassani</td>
</tr>
</tbody>
</table>